X-551-72-47

PREPRINT

NASA TM X 65847

ON RECONSTRUCTING TRAJECTORIES IN THE VENUS LOWER ATMOSPHERE

P. ARGENTIERO
G. WYATT

FEBRUARY 1972

GSFC — GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

X-551-72-47

ON RECONSTRUCTING TRAJECTORIES IN THE VENUS LOWER ATMOSPHERE

P. Argentiero
G. Wyatt

February 1972

Goddard Space Flight Center
Greenbelt, Maryland

CONTENTS

TABLES

ILLUSTRATIONS

ON RECONSTRUCTING TRAJECTORIES IN THE VENUS LOWER ATMOSPHERE

P. Argentiero
G. Wyatt

ABSTRACT

This paper utilizes a Monte Carlo technique in order to demonstrate the feasibility of processing in-situ measurements of temperature, pressure, and molecular weight in order to reconstruct trajectories in the Venus lower atmosphere. The technique assumes that the Venus lower atmosphere obeys the ideal gas law and the hydrostatic equation. Time correlations in the data are assumed to exist. It is also shown that the errors in trajectory reconstruction are due mostly to noise on the data rather than inaccuracies in the numerical technique.

ON RECONSTRUCTING TRAJECTORIES IN THE VENUS LOWER ATMOSPHERE

INTRODUCTION

The problem of reconstructing a trajectory in an atmosphere has been discussed in the literature. See for instance Preslin [1], Peterson [2], Sommer and Boissevain [3], and Seiff [4]. These studies assume that in situ measurements of acceleration, temperature and pressure are available. In the Venus lower atmosphere the density is so great that measurements from on board accelerations are not likely to be useful. Thus trajectory reconstruction in the Venus lower atmosphere must rely on range rate measurements and in-situ measurements of pressure, temperature, and molecular weight. The use of these data types in a standard parameter estimation mode requires that the planetary atmosphere be characterized by a relatively small number of parameters. This is usually done by assuming that the atmosphere satisfies the hydrostatic equation and the ideal gas law with the temperature profile piecewise linear with a small number of breakpoints. In this case the surface pressure and the temperature at the breakpoints along with the height of the breakpoints are parameters which define an atmosphere. A standard Kalman filter approach can then be implemented and an estimate of the trajectory can be obtained. The atmospheric parameters can be estimated as part of the state or they can be placed in a consider mode where they are not estimated but their associated uncertainties are permitted to have an influence on the filter behavior. The ensemble properties of such a filter can be studied in an error analysis mode and if the modeling assumptions are valid and if the non-linearities of the problem are not severe, the promulgated covariance matrix of such an error analysis represents an accurate measure of the statistical quality of the estimate.

Unfortunately there is no reason to believe that the non linearities of the problem are negligible. There are also present some obvious modeling errors in this approach. For instance it is doubtful if the temperature profile of the lower Venus atmosphere is approximately piecewise linear. The significance of such modeling errors is not now known. Consequently the results of an ensemble calculation associated with an error analysis of the parameter estimation technique is not to be trusted as an accurate indicator of the statistical quality of the technique. Another method for obtaining information concerning the statistical spread of the parameter estimation procedure is to perform a Monte Carlo study. But the computational algorithms for the implementation of the parameter estimation technique are too lengthy to permit such a possibility. One is forced to

conclude then that there is no effective way of determining the statistical quality of the standard parameter estimation technique as applied to trajectory determination in the Venus lower atmosphere.

There is a way out of this impasse. If one is willing to forego the direct use of range rate data, then a computational algorithm can be devised which utilizes temperature, pressure, and molecular weight data in order to estimate the trajectory. This algorithm is sufficiently simple that its ensemble properties can be accurately determined by means of a Monte Carlo study. This algorithm is obtained by assuming that measurements are unbiased and that the atmosphere obeys the hydrostatic equation and the ideal gas law. No assumptions concerning the temperature profile are necessary. This paper is a report on a Monte Carlo study of such an estimation procedure.

ESTIMATING ATMOSPHERIC TRAJECTORIES WITH PRESSURE, TEMPERATURE, AND MOLECULAR WEIGHT DATA

Assume that an atmosphere is perfectly mixed, and that it obeys the perfect gas law and the hydrostatic equation. Let P(H), T(H), ρ(H) be respectively the pressure, temperature and density of the atmosphere at a height H above the surface. Then

$$P(H) = \frac{\rho(H)\, R\, T(H)}{M} \tag{1}$$

and

$$\frac{dP(H)}{dH} = -g\,\rho(H) \tag{2}$$

where R is the universal gas constant, M is the molecular mass of the gas and g is the gravitational constant of the planet. (For present purposes the variation of g with H can be ignored.)

Assuming a ninety degree flight path angle, the trajectory of a probe in the Venus lower atmosphere can be described by a height versus time function H(t). Define the following functions

$$\begin{aligned} P(t) &= P[H(t)] && \text{(a)} \\ T(t) &= T[H(t)] && \text{(b)} \\ \rho(t) &= \rho[H(t)] && \text{(c)} \end{aligned} \tag{3}$$

Equations 1 and 2 yield

$$\frac{dH(t)}{dP(t)} = \frac{-RT(t)}{gP(t)M} \tag{4}$$

By multiplying both sides of equation 4 by DP(t)/Dt one obtains

$$\frac{dH(t)}{dt} = \frac{-R}{gM} \frac{T(t)}{P(t)} \frac{dP(t)}{dt} \tag{5}$$

and equation 5 yields

$$H(t) = \frac{R}{gM} \int_{t}^{t_0} \frac{T(t)}{P(t)} \frac{dP(t)}{dt} \, dt \tag{6}$$

where t_0 is the time of impact. Equation 6 provides an effective algorithm for reconstructing the height of the probe as a function of time.

The salient difficulty with the approach suggested by equation 6 is that the readings of temperature and pressure are available in discrete rather than continuous form. Thus both the derivative of pressure with respect to time and the integration of equation 6 must be obtained according to some numerical procedure. It is by no means obvious which numerical techniques would yield the smallest errors in the reconstruction of the trajectory. The difficulty is further compounded by the fact that there is noise on the data. Thus the error in trajectory reconstruction is ultimately the result of the interaction between the deterministic features of the numerical technique and the ensemble features of the data noise. The only systematic way to decide which of a set of numerical techniques yields the smallest errors in trajectory reconstruction is to test each in a Monte Carlo mode and to choose the numerical technique with the smallest critical error values as the most accurate. In the process one not only determines the best numerical technique but he also establishes the critical values of the error distribution of the resultant trajectory determination if the technique were utilized.

THE MONTE CARLO SIMULATION

In order to perform a Monte Carlo simulation, one must first construct a deterministic model of reality. In this case reality consists of a model of the Venus atmosphere and a probe with a certain trajectory which is sampling the

temperature, pressure, and molecular weight of the atmosphere according to a time schedule.

Since the atmosphere is assumed known and the trajectory and data sampling schedule are fixed, one knows the true values of the collected data. A random number generator is used to add noise to the true values of the data. This noise is so constructed as to follow a predetermined probability law. After the noise has been added, the resultant numbers are processed by the computational algorithm in use and an estimate of the trajectory is obtained. Since the true value of the trajectory is assumed known, one can determine the errors which the algorithm yields. The entire process can be repeated many times and an estimate of, say, the 95% critical value of the error in trajectory estimation at any given height can be obtained. The noise on the measurements was assumed to be zero mean and normally distributed. The assumed standard deviations as obtained from John Ainsworth of the Laboratory For Planetary Atmospheres, G.S.F.C., are displayed in Tables I and II.

Table I
Standard Deviation of Pressure Measurements

Range in Earth Atmosphere	Standard Deviation as Percent of Largest Value in Interval
90 - 20	.1%
20 - 4	.2%
4 - .08	.4%
.08 - .05	.5%

Table II
Standard Deviation of Temperature Measurements

Range in Kelvin	Standard Deviation as Percent of Length of Interval
800 - 600 K	1%
600 - 400 K	.75%
400 - 300 K	.60%
300 - 200 K	.5%

The G.S.F.C. 3609 model for the Venus atmosphere was used for our simulations. The probe trajectory was that of an entry probe with a 90° flight path angle and a parachute opening at 50 km above the surface. The reconstruction process is assumed to begin at 75 km above the surface. The trajectory was

constructed in the G.S.F.C. 3609 atmosphere by using the planetary explorer main probe configuration and assuming the quasi-static approximation. This approximation implies that the downward force of gravity is always in approximate equilibrium with the upward drag force. In effect this provides the velocity of the probe as a function of the density of atmosphere and makes for an easy trajectory reconstruction. Recently, this technique has been compared to a fourth order Runge Kutta technique and has been found to be valid in the Venus lower atmosphere, Reference [5].

Pressure, temperature, and molecular weight were assumed to be sampled a total of seventy-six times. The time intervals between measurements were constant and the first measurements were taken at 75 km and the last at the surface. Certain systematic errors in the measurement processes induce time correlations in the data. To model these measurements as independent is to make an optimistic error since this exaggerates the information content of the data. In this study the authors assume a positive correlation of .8 between a temperature or pressure measurement and the previous temperature or pressure measurement. These correlations become respectively .6, .4 and .2 when the time lag becomes 2, 3 and 4. When the time lag is 5 or greater we assume the measurements are independent. No cross correlation between the temperature and the pressure measurements are assumed to exist. The standard deviation of the noise on the measurements of molecular mass is assumed to be .5% of the nominal value of 43 as used in the 3609 model of the Venus atmosphere. A systematic error of .001 cm/sec^2 was assumed to have been made in the calculation of the surface gravitational acceleration of Venus.

The standard way to obtain Monte Carlo estimates of 95% critical values of errors at various heights is to execute the algorithm under investigation N times with the noise on the data chosen each time according to the probability law defined above. Each of the N trajectories so obtained can be compared with the true trajectory and at any given height, the 95% largest error from the ensemble of N errors is the optimal estimate of the 95% critical value of the error of the trajectory at that height. Unfortunately this optimal Monte Carlo technique when applied to this problem, leads to intractible problems both with regard to computer time and core storage. Thus a sub optimal Monte Carlo technique was utilized. The details of this technique are discussed in the appendix. This suboptimal technique does not sacrifice rigor and one can be sure to a confidence level of 99% that the estimate of a critical value it yields is larger than the correct critical value. In this sense the technique can be used to produce conservative estimates of critical values.

RESULTS AND CONCLUSIONS

Figures 1 through 4 provide graphs of respectively the 90%, 95%, 97% and 99% critical values of the error in trajectory reconstruction as a function of height above the surface. The algorithm used was implied by equation 6 with the derivative of pressure with respect to time calculated by assuming that pressure is exponential with time and with the integration implemented with the aid of a Simpson's rule algorithm. When the pressure derivatives are calculated by means of a simpler linear approximation, the errors increase by almost an order of magnitude. The use of a Simpson's rule in order to implement the integration instead of a simpler trapezoidal rule purchased an improvement in the errors of about 5%. Obviously the way the pressure derivatives are estimated is far more significant in the trajectory reconstruction than how the integration is performed.

From Figures 1 through 4 it is evident that at least to a height of 50 km above surface the processing of temperature and pressure data according to equation 6 yields a good estimate of the probe's trajectory. Above this height, the modeling assumptions, primarily the assumption that pressure measurements are exponential in time, become increasingly invalid. This above-mentioned exponential assumption rests on the assumptions that pressure is exponential with height and that probe velocity is constant. Clearly the greater the height above surface the greater is the acceleration the probe processes and the more questionable the assumption of constant velocity becomes. This is the primary impediment to the extension of this trajectory estimation technique to greater heights above surface. Of course, the technique's effectiveness could be extended considerably if the data acquisition rates were altered so that the higher the probe is above surface the greater the data acquisition rate. Although the authors have not done so, the possibility of extending this technique to much greater heights by means of a varied data acquisition rate could be systematically studied with the aid of the Monte Carlo techniques demonstrated in this report.

It is of some value to understand how much of the errors in trajectory reconstruction is due to noise on the data and how much is due to inaccuracies inherent in the numerical technique that one decides to use. In the present case such information is easy to obtain. Figure 5 provides a graph of the error obtained in trajectory reconstruction with perfect data. In this case the errors are due entirely to inaccuracies in the numerical technique. It is evident in comparing the information in Figure 5 with the information in any of the other figures that noise on the data rather than errors in the numerical technique is primarily responsible for errors in trajectory reconstruction. For instance compare the 95% critical error value in Figure 2 at 40 km, to the deterministic error at 40 km as given in Figure 5. The value from Figure 5 is smaller than the corresponding value from Figure 2 by an order of magnitude. The values of Figure 5

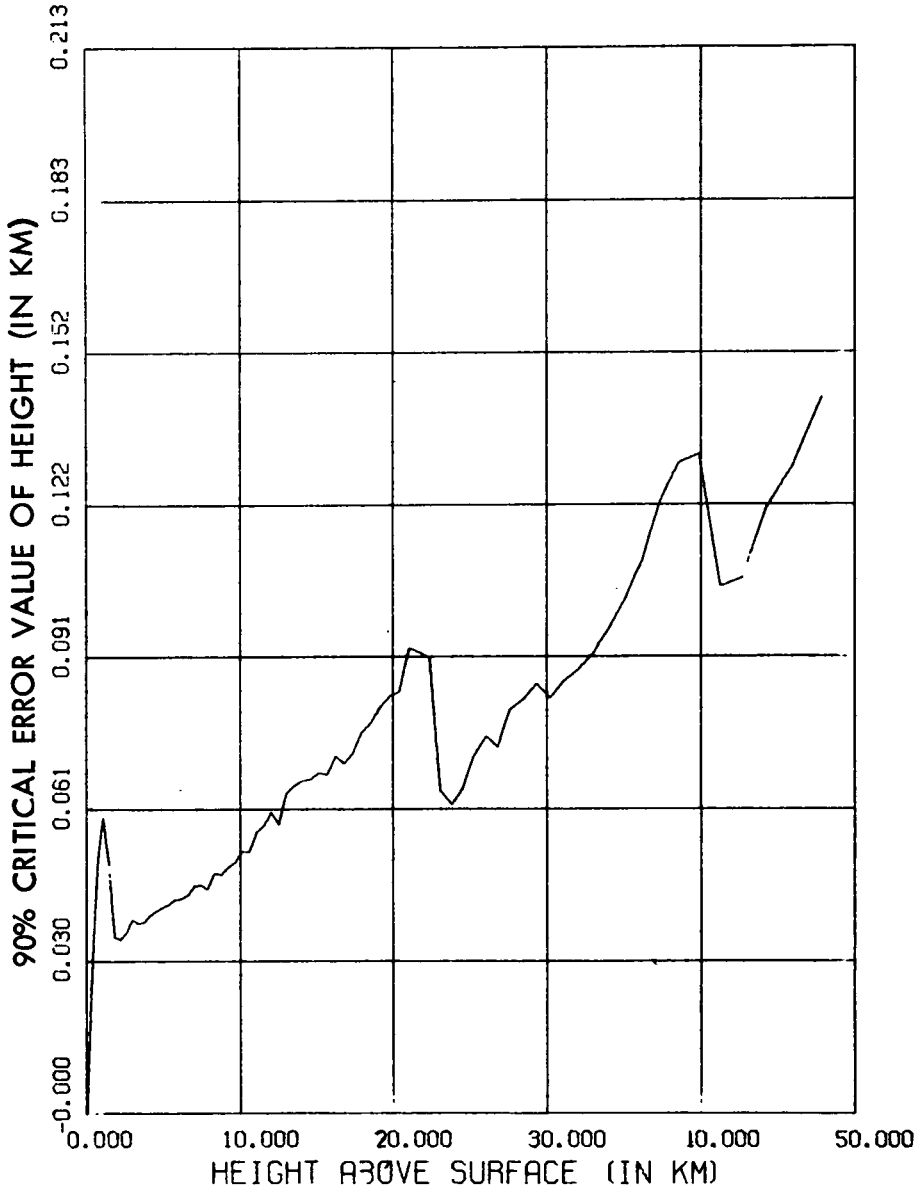

Figure 1. 90% Critical Error Value in Trajectory Reconstruction as a Function of Height Above Surface

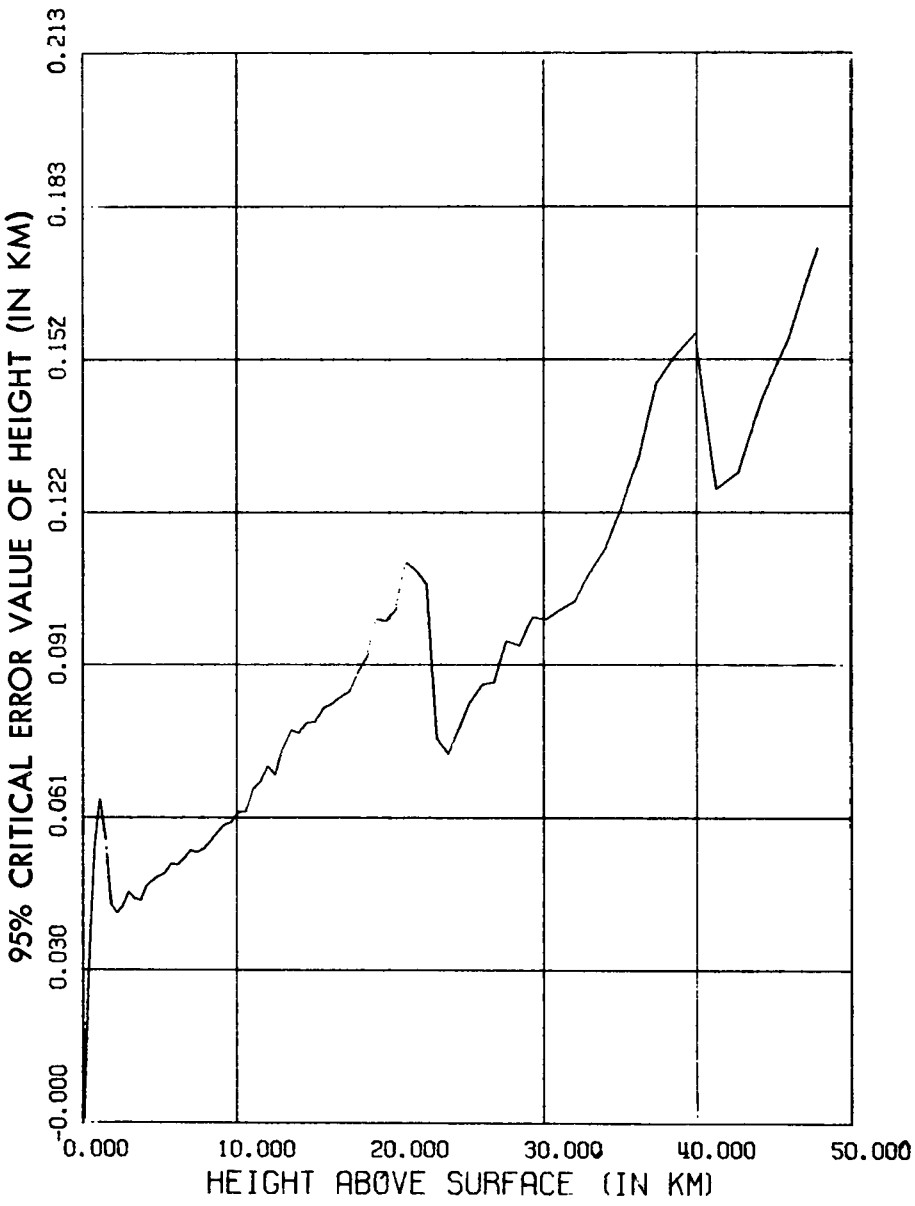

Figure 2. 95% Critical Error Value in Trajectory Reconstruction as a Function of Height Above Surface

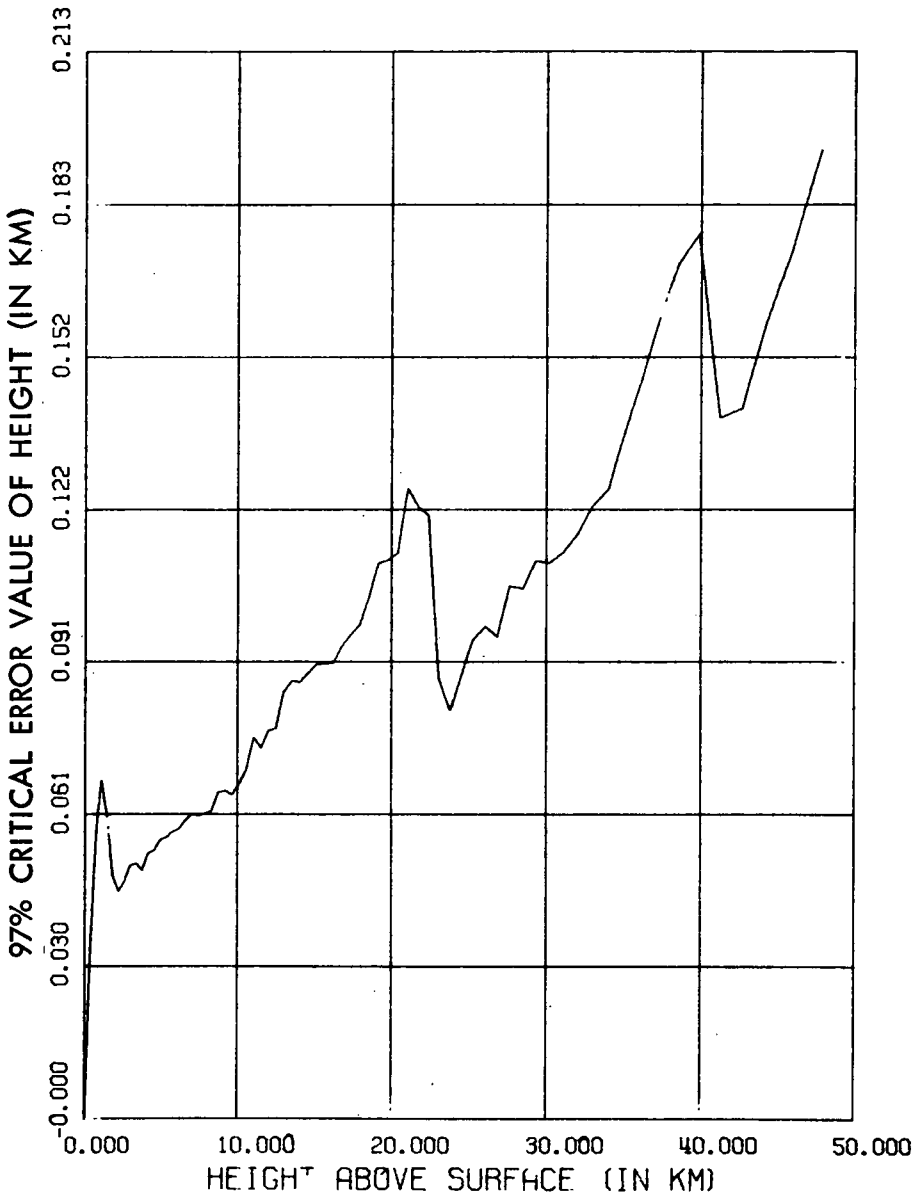

Figure 3. 97% Critical Error Value in Trajectory Reconstruction as a Function of Height Above Surface

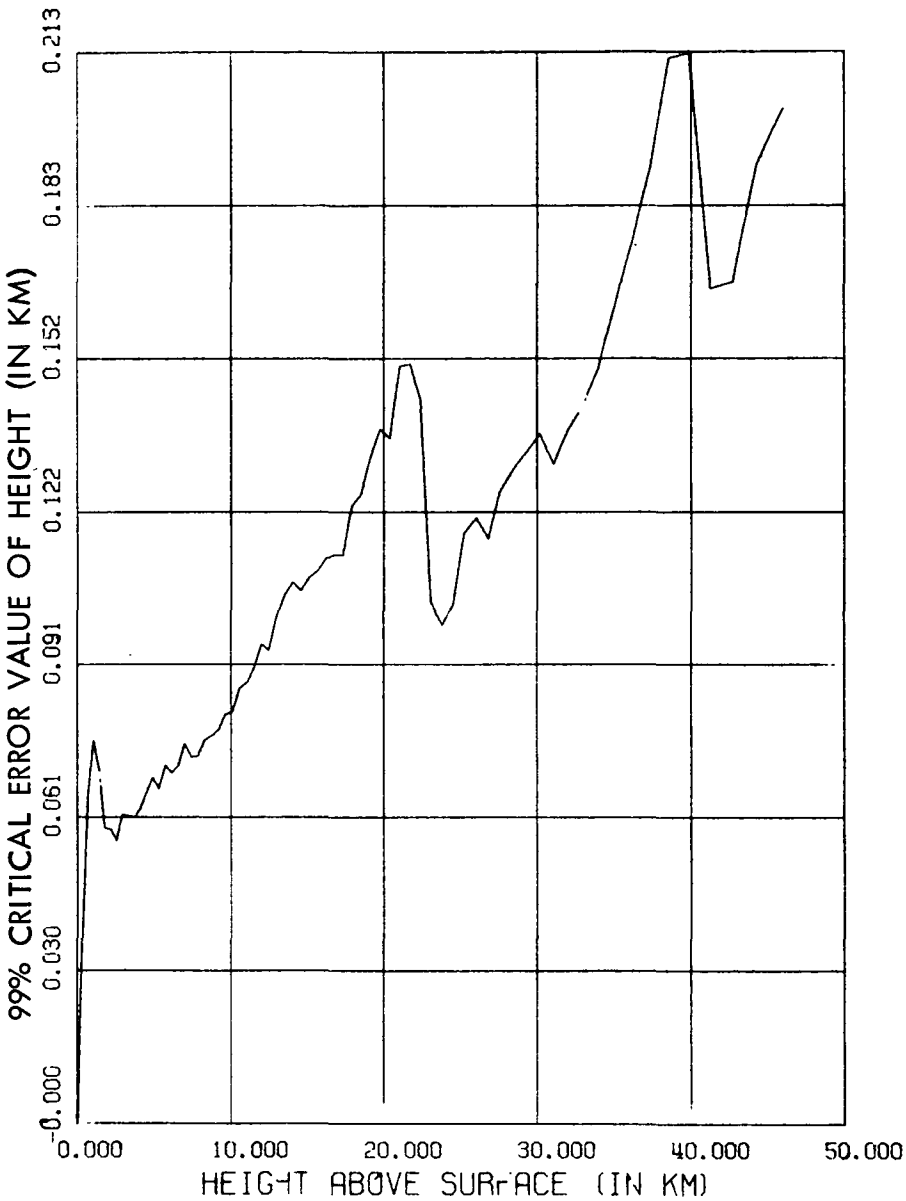

Figure 4. 99% Critical Error Value in Trajectory Reconstruction as a Function of Height Above Surface

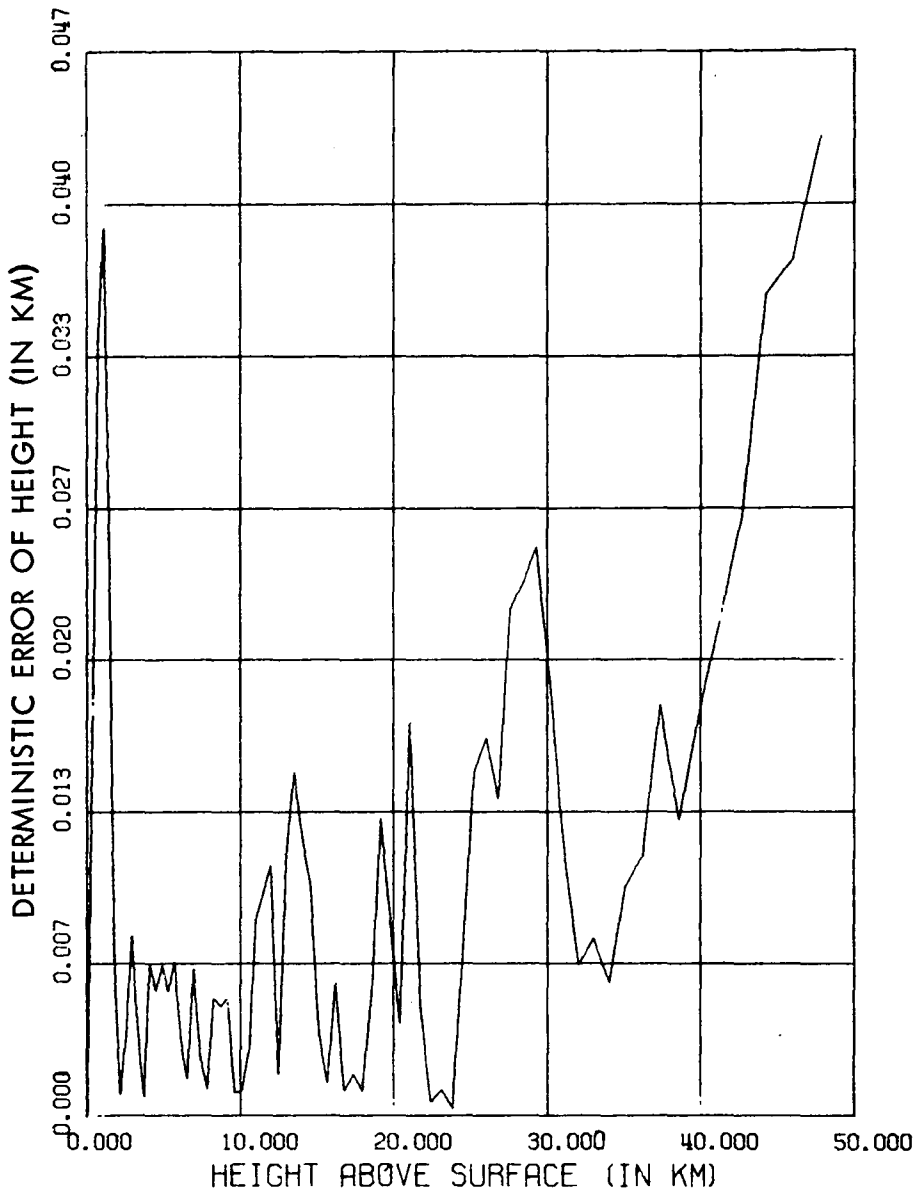

Figure 5. Deterministic Error Value in Trajectory Reconstruction as a Function of Height Above Surface

may be thought of as the values toward which the errors in trajectory reconstruction would asymptotically tend as the data became increasingly more accurate. As such these values represent the limit of how much can be purchased by improvement in data quality. Clearly there is substantial room for improvement.

REFERENCES

1. Prislin, R. H.; Errors Associated with Atmosphere Reconstruction from Capsule Onboard Measurements. JPL Space Programs Summary 37-46, Vol. IV, pp. 123-129.

2. Peterson, V. L.; Analysis of the Errors Associated with the Determination of Planetary Atmosphere Structure from Measured Accelerations of an Entry Vehicle, NASA TR R-225. National Aeronautics and Space Administration, Washington, D.C. July 1965.

3. Sommer, S. C., and Boissevain, A. G.; The Structure of an Atmosphere from Onboard Measurements of Pressure, Temperature, and Acceleration, NASA TN D-3933. National Aeronautics and Space Administration, Washington, D. C., April 1967.

4. Seiff, Alvin; Some Possibilities for Determining the Characteristics of the Atmospheres of Mars and Venus from Gas-Dynamic Behavior of a Probe Vehicle, NASA TN D-1770, 1963.

5. Analytical Manual for Landing Trajectory Reconstruction Computer Programs; prepared by Martin Marietta Corporation, Denver Division, for GSFC under Contract NAS5-11873.

APPENDIX

OPTIMAL AND SUB OPTIMAL MONTE CARLO TECHNIQUES FOR ESTIMATING CRITICAL VALUES

One usually employs a Monte Carlo technique in order to estimate a parameter associated with a probability distribution. In this study we have been concerned with the Monte Carlo estimation of critical values of a distribution. What follows below is an analysis of both an optimal and a sub optimal Monte Carlo technique for estimating critical values.

Given a sample of N independent valuations of a random variable, the "optimal" processing of this data in order to estimate the critical value of the random variable which corresponds to probability P can be very simply described. Let K be the smallest integer which is larger than NP. Linearly order the samples in descending order of magnitude. The K^{th} number in such a list is taken as the estimate of the critical value corresponding to probability P. We have yet to define in what sense such an estimate is optimal nor have we specified how the quality of the estimate depends on the sample size N.

In order to discuss these questions, it is first necessary to analyze the probability structure of the following experiment. Choose a probability P according to a uniform probability law. Then perform N independent Bernoulli trials with probability of success P. Divide the number of successes in N trials by N and call the result P'. The result of the experiment is the two tuple (P, P'). If N is sufficiently large, then an application of the normal approximation to the binomial distribution yields the conditional probability density function (p.d.f.) of P given that P' was fixed at some value α, $0 < \alpha < 1$ as

$$F(P) = \frac{1}{\sqrt{2\pi}\sqrt{\frac{P(1-P)}{N}}} \exp -\frac{1}{2}\left[\frac{P-\alpha}{\sqrt{\frac{P(1-P)}{N}}}\right]^2 \qquad (1)$$

Whenever one performs N Bernoulli trials where the probability of success is unknown, and the ratio of the number of successes to N is α, then equation 1 may be considered as an approximation to the p.d.f. of the true probability of success P.

With regard to the previously defined estimation procedure, the Bernoulli trials may be thought of as the independent sampling from the distribution and success in a trial can be defined as the probability of exceeding our estimated critical value. Since we choose this critical value such that the ratio of successes to trials is approximately α, equation 1 is the approximate probability density function of the true probability of the random variable in question exceeding our estimated critical value. The maximum value of the right side of equation 1 is achieved when $P = \alpha$. Thus the maximum likelihood estimate of the probability level associated with the estimated critical value is α. It is in this sense that the estimate is optimal.

We would like to obtain a region about the point $P = \alpha$ which would contain say 95% of the area under the curve defined by equation 1. If such regions could be constructed, then questions concerning the accuracy of the estimation technique could be adequately answered. To do this we substitute a probability density function whose integrated values are tabulated and which has the property that a 95% critical region about α contains more than 95% of the area under the p.d.f. defined byequation 1. A hint as to how this could be done is obtained by observing that the following inequality is valid for $0 < P < 1$

$$\frac{1}{2\sqrt{N}} \geq \sqrt{\frac{P(1-P)}{N}} \tag{2}$$

The substitution of the left side of inequality 2 for the expression on the right side transforms the right side of equation 1 into

$$g(P) = \frac{1}{\sqrt{2\pi}\ \frac{1}{2\sqrt{N}}} \exp -\frac{1}{2}\left[\frac{P-\alpha}{\frac{1}{2}\sqrt{N}}\right]^2 \tag{3}$$

Equation 3 provides the p.d.f. of a normal random variable with expectation α and standard deviation equal to $1/2\sqrt{N}$. In order to determine if this distribution has the necessary properties as a bound on the p.d.f. of equation 1, we must discover for what range of values $g(P)$ bounds $F(P)$. This is done by obtaining an expression in P for the log of the ratio of $g(P)$ to $F(P)$. The values of P for which this expression is positive determines the region for which $g(P)$ bounds $F(P)$. After considerable algebraic manipulation, the inequality to be solved takes the following form.

$$0 < P(1-P)\ \mathrm{LN}\ 2\sqrt{P(1-P)} + (P-\alpha)^2 N (4P^2 - 4P + 1) \tag{4}$$

It is readily shown that if P is in the region

$$\left\{P: \ |P - \alpha| \geq \frac{1}{\sqrt{N}}\right\} \tag{5}$$

then inequality 4 is satisfied. The area under g(P) in this region is .05. The area under F(P) in this region must be less than this value. Thus the p.d.f. of equation 3 has the desired properties as a bound.

It now becomes easy to determine the relationship between the accuracy of the estimation technique and the sample size N. If α is the estimated probability level associated with the critical value, then one is certain to a confidence level of .95 that the true probability level will not differ from α by more than $1/\sqrt{N}$. To be certain to this confidence level that the estimated probability level is not incorrect by more than .01, one needs a sample size of 10,000. If a confidence level of .99 is desired, then a sample size of 15,000 is needed.

The previous discussion has outlined the standard Monte Carlo technique for obtaining optimal estimates and has indicated the relationship between sample size and the accuracy. The technique has shortcomings, one practical and the other theoretical. To start with the practical difficulty, the technique can be demanding computationally if large sample sizes are involved. All the sample values must be stored in the computer at one time in order to implement any of the algorithms whose execution times are not prohibitive and in some circumstances this is not possible. The theoretical difficulty is that the optimal technique permits us to make probabilistic statements concerning the true probability level associated with our estimate of, say, the critical value associated with the .05 probability level. But we can say nothing probabilistically about the correct value of the critical value associated with the .05 probability level. For instance it could be useful in many circumstances to specify an interval and be certain to a given confidence level that the true critical value is in the interval. There exist so-called sub optimal Monte Carlo estimation techniques which can provide such information and which algorithmically are more tractible than the optimal technique. One such technique, the one that was utilized to obtain the results quoted in the body of this report will now be described.

Suppose one obtains K batches of samples each containing N values and all the values are assumed to be independent. For each of the K batches we can process N samples in the batch in the optimal fashion described previously and thus obtain K numbers X_i, $i \leq K$ all of which are estimates of the critical value X associated with a probability level α. Each of these estimates is equally likely to be less than as greater than the correct critical value X. Thus the

number of our K estimates which exceed X is a binomial random variable with expectation K/2 and variance K/4. Using the usual normal approximation to the binomial distribution we can infer to a confidence level of .99, that X is smaller than at least M_2 of the K values where M_2 is the largest integer which is smaller than $K/2 + 3\sqrt{K}/2$ and that X is larger than M_1 of the values where M_1 is the smallest integer which is larger than $K/2 - 3\sqrt{K}/2$. The significance of these facts becomes apparent when we state them in the following logically equivalent form. Suppose we order the K estimates of X in ascending order of magnitude. Then one is sure to a confidence level of .99 that the correct critical value X lies between the value in the sequence indexed by M_1 and the value in the sequence indexed by M_2. Thus by this technique we have succeeded in structuring an interval in which we are certain to a confidence level of .99 the correct critical value X resides. The algorithmic advantages of this technique rest chiefly in the fact that only one batch of values at a time need be stored in the computer. There are also advantages in terms of computing speed.

While rigorous statistical results concerning the accuracy of this technique are difficult to obtain, it is intuitively evident how the batch size N and the number of batches K influence the accuracy. The number N determines the accuracy of each individual estimate, that is, N determines how widely scattered the estimates are about the true value X and thus how large the interval which is expected to contain X will be. If N is small these estimates will be widely scattered and the interval will be much larger than necessary and thus too conservative. The impact of the number of batches K on the accuracy of the technique is more subtle but nevertheless quite significant. First, if K is too small the normal approximation to the binomial distribution becomes invalid and serious errors can result. In addition to this factor there is the fact that we are not at liberty to choose any confidence level we wish since there are only a finite number of intervals that this technique permits one to choose. We pick the interval whose confidence level bounds most closely the desired confidence level. The larger the value of K which is chosen, the better the quality of this bound.

NASA–GSFC COML., Arlington, Va.